

On the areal dimension of non-standard syntax: evidence from a Portuguese dialect corpus

Ernestina Carrilho and Sandra Pereira

University of Lisbon

Introduction¹

This paper investigates aspects of the areal dimension of dialect syntax. More concretely, the investigation focuses on the identification of syntactic areas within the Portuguese territory and on their connection with traditional dialect-geographic areas of European Portuguese (henceforth EP). After some brief methodological considerations, the main part of the paper presents the geographical distribution of a selection of non-standard syntactic constructions in EP; the final part further illustrates how syntactic areas compare to dialectal areas defined on the basis of other type of linguistic data (namely, phonetic data). Dialect syntax is considered here from a geolinguistic perspective aimed at: (i) showing how non-standard syntax may be geographically represented; (ii) relating such syntactic areas to more general dialectal areas. As a broad conclusion, the paper highlights the contribution of dialect syntax for geolinguistic purposes.

Methodological background

Although the field of syntax can hardly be regarded as a prolific one in dialectology, the last two decades have given rise to important developments in the study of dialect syntax.² Empirical, methodological and conceptual advances in this recent field encourage enhanced approaches to the spatial dimension of

¹ This work was supported by Fundação para a Ciência e a Tecnologia (FCT) - grant SFRH/BD/27648/2006 (S. Pereira) and project PTDC/LIN/71559/2006. We would like to thank the editors of this volume, for their comments and suggestions, and the organizers and participants of the *Workshop Dialect and Regiolect Syntax*, held within *Methods in Dialectology 14* (University of Western Ontario, August 2011), for helpful discussion.

² Detailed information on recent and current research in dialect syntax has been made available by the *Edisyn* project at <http://www.dialectsyntax.org>.

linguistic variation, which may take advantage of new insights into the areal study of syntactic constructions.

The approach developed in this paper relies on the empirical exploitation of a dialect corpus, the *Syntax-oriented Corpus of Portuguese Dialects*, CORDIAL-SIN.³ Over the past decade, studies on different aspects of EP dialect syntax have been carried out in tandem with the compilation of this corpus.⁴ The geolinguistic expression of syntactic variation through CORDIAL-SIN is under consideration in this paper (as also in Carrilho & Pereira 2011; Carrilho & Lobo 2012).

CORDIAL-SIN is an annotated corpus compiling spontaneous or almost spontaneous speech drawn from tape-recorded dialect interviews that were conducted for different linguistic atlases.⁵ The corpus amounts to more than 600,000 words, from 42 locations distributed over the continental and insular Portuguese territory. These data are representative of traditional dialects, as spoken by informants who are aged, non-educated, rural and born and raised in the place of interview.

The areal distribution of EP non-standard constructions

Differently from other non-standard constructions which spread over all dialects (even if absent from standard EP), the constructions under inspection in this paper mainly occur in restricted areas in the Portuguese territory. The following sections present the geographical distribution of a selection of non-standard constructions involving: (i) inflected gerunds; (ii) periphrastic aspectual forms with *estar* ‘to be’ followed by a gerund; (iii) third person plural verbal agreement with the pronominal form *a gente* (literally ‘the people’, third person singular); (iv) verb *ter* ‘to have’ in impersonal existential constructions; and (v) the absence of a definite article before pronominal possessives.

³ CORDIAL-SIN is completely downloadable from <http://www.clul.ul.pt/en/resources/411-cordial-corpus> and is also searchable through the *Edisyn Search Engine* at <http://www.meertens.knaw.nl/edisyn/searchengine/>.

⁴ Most of these studies are available through <http://www.clul.ul.pt/en/resources/212-cordial-sin-syntax-oriented-corpus-of-portuguese-dialects>, where more detailed information on CORDIAL-SIN may also be found (namely the identification of the CORDIAL-SIN locations represented by a code in the examples in this study, such as LVR05 in example (1)).

⁵ These dialectal interviews were gathered by the Dialectology Group of the Centro de Linguística da Universidade de Lisboa, within the scope of the atlases ALEAç, ALEPG, ALLP.

Inflected gerund

Some EP varieties display the presence of person inflection in gerund forms (examples 1a, 2a), while the gerund is categorically invariable in standard EP (examples 1b, 2b):

- (1) a. Em eles **tendem** a outra, já não querem aquela. (LVR05)
 in they have.GER.3PL the other.one already not want.3PL that.one
 ‘when they have the other one, they don’t want that one anymore’
- b. **Tendo** eles a outra, ...
 have.GER they the other.one
- (2) a. E **tendem** uma árvore, não há pássaro nenhum que poise no chão. (STJ30)
 and have.GER.3PL a tree not has bird any that land on.the ground
 ‘and if a bird has a tree, it does not land on the ground’
- b. E tendo uma árvore, ...
 and have.GER a tree

Monographic studies of EP dialects have often noticed instances of inflected gerunds, which, according to Lobo (2001, 2008), mainly occur in adverbial clauses, as in the examples above. The areal distribution of these inflected forms in CORDIAL-SIN has been investigated in Lobo (2008): as represented in Map 1, the non-standard variant mainly appears in an homogeneous area in the southern part of Portugal and in one location in the insular territory (in Azores).⁶

⁶ The inspection of this non-standard variant faces particular difficulties, as noticed by Lobo (2008): since the inflected forms for first person singular and third person singular coincide with the non-inflected gerund form (for instance, *tendo* in 1b), the absence of data may in some cases mean the absence of unambiguous occurrences of the inflected gerund (which may be the case in the blank spots found in the southern area, for which inflected gerunds have been observed in monographic studies).



Map 1: Distribution of inflected gerunds in CORDIAL-SIN (adapted from Lobo 2008)

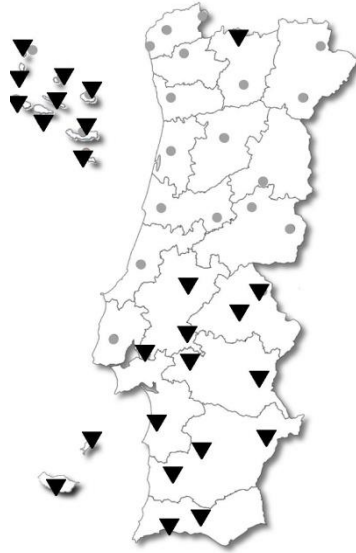
Periphrastic aspectual construction *estar* + gerund

As a variant of the standard periphrastic aspectual construction illustrated by 3a and 4a (which involve inflected forms of *estar* ‘to be’ followed by the aspectual preposition *a* and the infinitival form of a main verb), some EP varieties display a periphrastic aspectual construction in which inflected *estar* combines with a gerund form of the main verb (examples 3b and 4b):⁷

- (3) a. Estão **a arder**
 are_{3PL} a.ASP burn.INF
 ‘They are burning’
- b. Estão **ardendo** (PAL36)
 are._{3PL} burn.GER
- (4) a. Essa pessoa estava **a varrer**
 that person was._{3SG} a.ASP sweep.INF
 ‘That person was sweeping’
- b. Essa pessoa estava **varrendo** (LUZ08)
 that person was._{3SG} sweep.GER

⁷ Besides the EP varieties under inspection, the gerund periphrastic construction is extensively found in Brazilian Portuguese. The construction involving *estar* is taken in this study as representative of the corresponding aspectual construction with other verbs, such as *ficar* ‘to remain’, *continuar* ‘to continue’, by far less frequent in CORDIAL-SIN.

Through the CORDIAL-SIN data, the gerund variant appears consistently spread over the southern part of Portugal and in the islands (Map 2).⁸



Map 2: Distribution of aspectual *estar* + *gerund* in CORDIAL-SIN

The pronominal form *a gente* agreeing with 3PL verbal forms

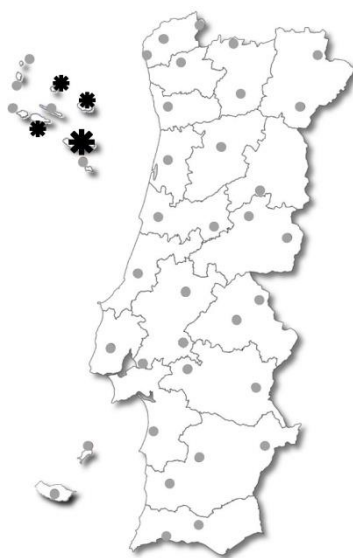
Variation in verbal agreement with the pronominal subject *a gente* is frequent in EP varieties (see Pereira 2003, for a recent survey). The pronominal form stems from a third singular phrase (literally ‘the people’) but has plural human reference including the speaker. This frequently correlates, in popular varieties, with occurrences of first person plural verb agreement with the subject *a gente*, while the standard variant displays third person singular agreement (examples 5a-b below, respectively). As CORDIAL-SIN data show, Azorean dialects further display a different pattern of agreement, involving third person plural, as illustrated in example 5c (Pereira 2003).

- (5) a. **A gente** não dávamos nome nenhum àquilo.
 the people.3SG not gave.1PL name any to-that
 ‘We didn’t give any name to that.’
- b. **A gente** não dava nome nenhum àquilo.
 the people.3SG not gave.3SG name any to-that

⁸ The presence of this variant at an isolated location on the northern border should be understood in a wider geolinguistic context, taking into consideration the distribution of this variant also in the Galician area (see Álvarez & Xove 2002: 361).

- c. **A gente** não davam nome nenhum àquilo. (MIG08)
the people.3SG not gave.3PL name any to-that

This regional variant is restricted to Azorean dialects. Among these, it appears with special frequency on the island of São Miguel, and only sporadically in some other locations, as shown differentially in Map 3.



Map 3: Distribution of *a gente* + 3PL verbal agreement in CORDIAL-SIN

Impersonal construction with existential verb *ter* ‘to have’

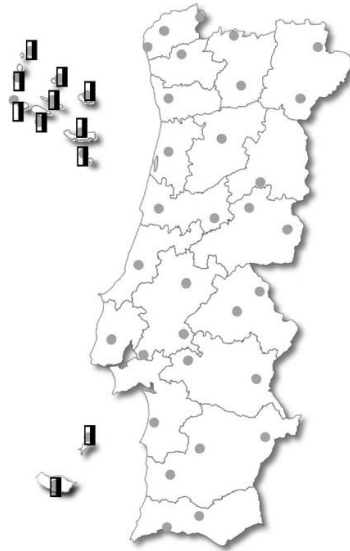
In standard EP, impersonal existential constructions involve a third singular form of the verb *haver* (etymologically ‘to have’, from Latin HABERE), as illustrated in 6a. The CORDIAL-SIN data also display semantically existential instances of the verb that expresses possession in standard EP, *ter* (‘to have’, from Latin TENERE).⁹ Thus, some EP varieties allow this existential use of *ter*, which occurs in impersonal constructions (regardless of the number displayed by the co-occurring constituent, the verbal form invariably displays third person singular agreement, as in example 6b below).

- (6) a. Mas **havia** muitos moinhos por aqui fora.
 but “had”.3SG many water-mills.PL by here out
 ‘But there were many water-mills this way out.’

⁹ This existential construction is commonly used in Brazilian Portuguese (Avelar 2004, among others).

- b. Mas **tinha** muitos moinhos por aqui fora. (CLH03)
 but had._{3SG} many water-mills._{PL} by here out

In CORDIAL-SIN, the non-standard construction discussed here is mainly found in the insular territory, both in Azores and in Madeira (Map 4).



Map 4: Distribution of impersonal existential *ter* in CORDIAL-SIN

Absence of article before pronominal possessives

The last examples concern the variable presence of the article before a pronominal possessive, a phenomenon often noticed as a matter of cross-linguistic and intra-linguistic variation (Mattos e Silva 1989; Rinke 2010; among others). The variant in question corresponds strictly to contexts in which the standard variety displays the article, as in examples 7a and 8a. In this case, the geographical spread of the non-standard variant (illustrated in 7b and 8b) is not homogeneous, as presented more extensively in Carrilho and Pereira (2011).

- (7) a. E **a** **minha** **tia** lia aqueles livros.
 and the._{FEM.SG} my._{FEM.SG} aunt read those books
 ‘And my aunt read those books.’
- b. E **minha** **tia** lia aqueles livros. (GRJ06)
 and my._{FEM.SG} aunt read those books

- (8) a. Olha, fala com **o teu avô**
 look, talk with the.SG your.SG grandfather
 ‘Look, talk to your grandfather.’
- b. Olha, fala com **teu avô** [...] (CTL18)
 look talk with your.SG grandfather

An important areal asymmetry emerges, however, which makes us consider this phenomenon among those non-standard constructions that are specific to certain EP areas. When one takes into account the relative frequency of the article before pronominal possessives (more accurately, before possessives preceding kinship nouns, as in the examples above), an evident geographical pattern emerges: while the non-standard construction occasionally occurs from the northern to the southern area of continental EP,¹⁰ only in some insular areas its incidence amounts to more than 65%, namely in the Madeira archipelago. The corresponding locations are represented in Map 5.



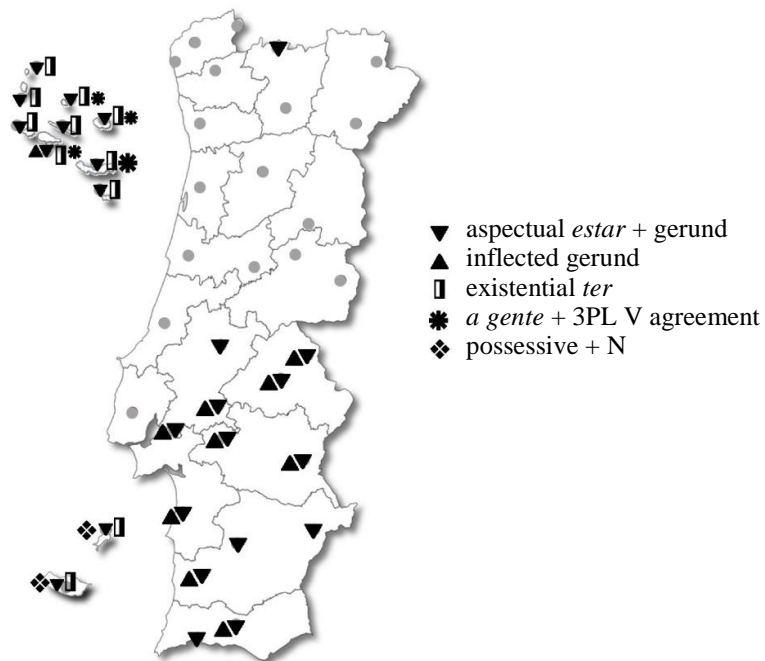
Map 5: Distribution of absence of article before pronominal possessive (> 65%)

¹⁰ This geographical pattern (evocative of “leopard spots”) has been noticed as fairly recurrent in the case of syntactic variants (Poletto 2010).

Syntactic areas and EP dialects

The non-standard constructions considered in the previous section contribute to the definition of specific syntactic areas in EP. For each non-standard variant a particular geographical distribution has been identified through CORDIAL-SIN. Although corpus data may, in certain cases, offer meagre evidence for constructions which are less frequent in a language, the constructions under study here could be systematically found throughout this EP dialect corpus. Moreover, the maps previously presented show that the occurrence of each non-standard variant is geographically well delimited: through Map 1 to Map 4, the blank areas represent homogeneous absence of the non-standard variant, while manifestations of this variant only appear and concentrate on the marked areas. As mentioned above, Map 5 differs somewhat in the type of data represented. Nevertheless, even in this case, a clear geographic asymmetry arises when one takes into account the higher preference for the non-standard variant.

The homogeneous distribution of the non-standard variants may thus straightforwardly lead to the identification of isoglosses for the phenomena at stake. Additionally, although the phenomena examined here display different geographical distributions, it is worth remarking that all the syntactic areas delimited by these isoglosses are concentrated over the southern part of the continental territory and over the Portuguese islands, as represented in Map 6.



Map 6: Summary of distribution of EP non-standard constructions in CORDIAL-SIN

Significantly, the geographical distribution of these non-standard constructions corresponds to important dialectal boundaries in EP, which have been identified on the basis of regional phonetic variants. Thus, if one considers the boundaries within EP continental dialects, as proposed by Cintra (1971), it becomes clear that the phenomena here inspected are limited, in the continental territory, to the area of Central Inland and Southern Portuguese dialects (whose Northern limit roughly follows the course of River Tagus in Portugal). In addition, some of these phenomena, in particular those involving the gerund, are equally found in insular dialects, which reinforces the linguistic similarity between both areas, as previously noticed on the basis of phonetic variants (Cintra 1971).

Although Cintra's proposal did not cover the insular territory, the identification of insular dialects is often acknowledged, and important differences have been noticed for dialects from Azores and dialects from Madeira and even within each of these groups of dialects (Cintra 1990; Segura 2006; among others). Importantly, some of the regional syntactic constructions inspected in this paper display an insular distribution which varies from generalized in both archipelagos (the non-standard existential construction) or restricted to one of them (the absence of article before possessive in Madeira) to particular in one insular area whose dialectal identity has independently been noticed (the case of third plural agreement with *a gente* in S. Miguel).

Summing up, it is possible to identify particular syntactic areas in the Portuguese territory, which has been accomplished in this paper by inspection of the geographical distribution of a selection of non-standard constructions in EP, as they occur through the dialect corpus CORDIAL-SIN. Taking into consideration the particular areas thus identified, it must be observed that syntactic variation strengthens, in these cases, important EP dialectal limits which have been established on the basis of phonetic differentiation, namely: (i) Central Inland and Southern dialects behave fairly differently from Central Littoral dialects and Northern dialects in the continental territory; (ii) Central and Southern dialects behave fairly similarly to insular dialects; (iii) insular dialects share some dialectal similarities but differ with respect to other dialectal aspects; (iv) inside the main dialectal areas, some syntactic variants help to identify local varieties.

Conclusion

On methodological grounds, the work presented in this paper emphasizes the relevance of dialect corpora as tools for the geolinguistic study of dialect syntax. The identification of syntactic areas in EP and the type of connection that these

areas establish with EP dialects contribute to a twofold general conclusion: (i) syntactic variation may be confined to cohesive dialect areas; (ii) syntactic areas provide additional geolinguistic support to the tasks of identifying and delimiting dialects.

References

- ALEAç – *Atlas Linguístico e Etnográfico dos Açores* (João Saramago, coord.) [http://www.clul.ul.pt/pt/investigacao/206-project-linguistic-and-ethnographic-atlas-of-azores-aleac]
- ALEPG – *Atlas Linguístico-Etnográfico de Portugal e da Galiza* (João Saramago, coord.) [http://www.clul.ul.pt/pt/investigacao/205-linguistic-and-ethnographic-atlas-of-portugal-and-galicia-alepg]
- ALLP – *Atlas Linguístico do Litoral Português* (Gabriela Vitorino, coord.) [http://www.clul.ul.pt/pt/investigacao/207-project-linguistic-atlas-of-the-portuguese-coast-allp]
- Álvarez, Rosario and Xosé Xove. 2002. *Gramática da Lingua Galega*. Vigo: Galaxia.
- Avelar, Juanito. 2004. *Dinâmicas morfosintáticas com ‘ter’, ‘ser’ e ‘estar’ em português brasileiro*. Master’s Dissertation. IEL-UNICAMP.
- Carrilho, Ernestina and Sandra Pereira. 2011. Sobre a distribuição geográfica de construções sintáticas não-padrão em português europeu. *Textos Seleccionados do XXVI Encontro Nacional da Associação Portuguesa de Linguística*. CD-ROM. Lisboa: APL.
- Carrilho, Ernestina and Maria Lobo. 2012. Contribution à l’étude de la variation syntaxique dans le domaine ibéro-roman. In Michèle Oliviéri, Guylaine Brun-Trigaud and Philippe Del Giudice (eds.) *La Leçon des dialects. Hommages à Jean-Philippe Dalbera*. Alessandria: Edizioni dell’Orso, 323-336.
- Cintra, Luís Filipe Lindley. 1971. Nova proposta de classificação dos dialectos galego-portugueses. *Boletim de Filologia* 22, 81-116.
- Cintra, Luis Filipe Lindley. 1990. Os dialectos da ilha da Madeira no quadro dos dialectos galego-portugueses. Published in José Eduardo Franco (ed.) (2008) *Cultura Madeirense. Temas e Problemas*. Porto: Campo das Letras, 95-107.
- CORDIAL-SIN – *Corpus Dialectal para o Estudo da Sintaxe / The Syntax-oriented Corpus of Portuguese Dialects* (Ana Maria Martins, coord.) [http://www.clul.ul.pt/en/research-teams/212-cordial-sin-syntax-oriented-corpus-of-portuguese-dialects]
- EDISYN – *European Dialect Syntax Project* (Sjef Barbiers, coord.) [http://www.dialectsyntax.org/index.php/project-description-edisyn-mainmenu-50]
- Lobo, Maria. 2001. On gerund clauses of Portuguese dialects. In Alexandre Veiga, Víctor Manuel Longa and J. Anderson (eds.) *El Verbo. Entre el Léxico y la Gramática*. Lugo: Ed. Tris Tram, 107-118.
- Lobo, Maria. 2008. Variação morfo-sintáctica em dialectos do Português europeu: o gerúndio flexionado. *Diacrítica* 22 (1). Braga: Universidade do Minho, pp. 25-55.
- Mattos e Silva, Rosa Virgínia. 1989. *Estruturas trecentistas: Elementos para uma gramática do Português Arcaico*. Lisboa: Imprensa Nacional – Casa da Moeda.
- Pereira, Sandra. 2003. *Gramática comparada de a gente – variação no PE*. Master’s Dissertation, Faculdade de Letras da Universidade de Lisboa.
- Poletto, Cecilia. 2010. Contrastive linguistics and microvariation. Communication at *6th International Contrastive Linguistics Conference*, Freie Universität Berlin, October 2010.
- Rinke, Esther. 2010. A combinação do artigo definido e pronome possessivo na história do português. *Estudos Linguística Galega* 2, 121-139.

Segura, Luisa. 2006. Dialectos açorianos. Contributos para a sua classificação. In Maria Clara R. Bernardo and Helena M. Montenegro (eds.) *I Encontro de Estudos Dialectológicos – Actas* (2003). Ponta Delgada: Instituto Cultural de Ponta, 325-344.